

## 『日本語日常会話コーパス』CEJC の調整頻度算出に用いる語数表の準備

## 『日本語日常会話コーパス』CEJC

<https://www2.ninjal.ac.jp/conversation/cejc.html>

モニター公開版の詳細については以下をご覧ください。

データの仕様・関連文献	関連データ・ツール	利用ガイドライン
データ概要	調査協力者一覧	利用者の範囲・利用期限
会話・話者のメタ情報	会話一覧	成果の公表に際して
転記テキスト	語彙表	利用目的の範囲・利用方法
短単位情報	その他のデータ	授業や演習等での利用について
映像・音声データ	ツール	バグ報告
関連文献		FAQ

「語彙表」をクリックすると下記ページへ ↓

①

ホーム » 日常会話コーパス » モニター公開 » 語彙表

## 語彙表・語数表

『日本語日常会話コーパス』モニター版の「中納言」データの語彙表および語数表 (ver.2020.03) を公開します。研究、教育目的であれば無償で自由にお使いになれます。

語彙表・語数表データ

- 語彙表と語数表の解説(PDF、以下ファイルの説明)
  - 0\_CEJC語彙表と語数表の解説\_ver202003.pdf
- 語彙表\_品詞構成表\_語種構成表(tsv形式) (Version 202003)のダウンロード
  - CEJC語彙表\_品詞構成表\_語種構成表(tsv形式).zip
- CEJC語彙表(xlsx形式) (Version 202003)のダウンロード
  - CEJC語彙表(xlsx形式)\_ver202003.zip
- CEJC語数表\_品詞構成表\_語種構成表(xlsx形式) (Version 202003)のダウンロード
  - CEJC語数表\_品詞構成表\_語種構成表(xlsx形式)\_ver202003.zip

「CEJC 語数表\_品詞構成表\_語種構成表(xlsx形式)\_ver202003.zip」をダウンロードする。

② zip ファイルを「すべて展開」する。以下のファイルがある。

名前

- 7\_cejc\_wc.xlsx
- 8\_CEJC\_語数表\_形式と場面と年齢と性別.xlsx
- 10\_cejc\_frequencylist\_pos.xlsx
- 12\_cejc\_frequencylist\_wtype.xlsx

※このうちの、「8\_CEJC\_語数表\_形式と場面と年齢と性別.xlsx」が、粗頻度から調整頻度を算出する場合に必要な語数表。